,,

Universiteit Leiden
The Netherlands

Informatica en Economie

The influence of behavioral richness on
human's perception of reception robots

Leendert Pieter van der Plas

Supervisors:
Joost Broekens & Walter Kosters

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

14/07/2020

Abstract

The current focus on “Cognitive Assistants”, a form of artificial intelligence that operates on user command, seems to be on virtual products instead of physical products. Products such as Google Home and Amazon Alexa have sold hundreds of millions of units while humanoid robots such as a Nao or Pepper robot have only sold tens of thousands of units.

As there is a vast price difference between the two types of products, this study will investigate if the extra functionality that a human-shaped robot adds, has a large enough effect on user experience. This is done by focusing on a Pepper robot in a hospitality position, by having this robot function as an addition to the reception personnel.

This study aims to determine the benefits of various conditions of interaction that are not possible on a purely virtual product. The measurement itself will be done through the Godspeed Questionnaire combined with some questions about usability at the end of the experiment.

Based on a review of literature in the Human robot interaction (HRI) field, there is little research about end user experience itself and the studies that have been done show that there is a substantial “Wow factor” effect that influences peoples opinion on the functionality of the robot. It is also unclear what different conditions of functionality add to the user experience.

Due to technical issues and the Corona virus limiting the amount of real life participants, videos showing how a participant would interact with the robot were used. Based on these videos a participant could then fill in the questionnaire. Due to this, the final condition of the experiment could not be evaluated.

The results show that expanding the testing conditions by adding more functionality does not lead to a statistically significant difference in any of these conditions. It is possible that there is no effect between these conditions but that there may be an effect if for example the fourth condition, which introduces pointing at and looking in a direction, is properly tested.

On this basis, it is not possible to conclude that there is any statistically significant difference in user experience and usability based on the three tested conditions, which introduce visually and/or verbally supported directions. Further research could either repeat this experiment to remove this video element and get more data or expand the scope of this experiment with more conditions, to see if that does result in a statistically significant difference.

Contents

1 Introduction

The current focus on "Cognitive Assistants", a form of artificial intelligence that operates on user command (Herron, 2017), seems to be on virtual products instead of physical products. Amazon's Echo, Apple's Siri, Google's Home, IBM's Watson and Microsoft's Cortana are all primarily virtual assistants, requiring only a device able to run the software itself. This is in contrast to SoftBank Robotics's approach with either the Pepper or the Nao, both human-shaped robots (Ebling, 2016). This sentiment is also evident in sales, where the Pepper robot, as of May 2018, has been sold 12,000 times in Europe (Olson, 2018). Amazon's Echo has sold over 100 million devices as of January 2019 (Matney, 2019).

A large part of this difference can be attributed to the price difference, as an Echo can be bought for 50 dollars, while a Pepper can be bought for 1,600 dollars. The required subscription of $360 per month, over 36 months, brings the total cost of ownership to over $14,000. Another factor could be the production of said devices, as when the Pepper robot was launched in June 2015, it sold out in 60 seconds (Mogg, 2015). This suggests that the demand is there, but cannot be met. Another large factor is that, currently, the Pepper robot is focused towards businesses instead of consumer homes. For example, the first commerce application for SoftBank Robotics's humanoid robot Pepper is a hospitality kiosk for Pizza Hut restaurants. (Ly, 2016)

One consumer use case of the Pepper robot could be companionship and elderly care. With its humanoid shape and its ability to complexly communicate with people, Pepper robots are particularly suited for this task (Lagrue & Busy, 2020). Various kinds of robots are already used in nursing homes in Japan, where these robots have to help fill a projected shortfall of 380,000 workers by 2025. A resident at one of these homes has stated that "These robots are wonderful", and that "More people live alone these days, and a robot can be a conversation partner for them. It will make life more fun.". While these robots are currently used in nursing homes, this is not the only place where people often live alone, suggesting that these robots can also provide companionship in other areas. The current global market of nursing care and disabled aid robots is still small, with a revenue of 19.2 million dollars in 2016, but is estimated to increase to 3.8 billion dollars by 2035 in Japan alone (Foster, 2018).

In this paper, we want to investigate the effect of the various extra conditions of functionality, that a human-shaped robot adds, on utility and user experience. This is done by focusing on the Pepper robot in a hospitality position, by having this robot function as an addition to the reception personnel.

1.1 Motivation

There are few studies to-date that investigated consumer interaction and experience with hospitality robots in real-world applications. The primary focus of engineering and computer science concerning hospitality robots is still on the technical aspects of such a device, the architecture and its performance, rather than on these consumer experiences despite this being one of the most important aspects of a hospitality or tourism robot. This paper will investigate consumer experience in the real world by looking at the features and tasks that a human-shaped robot adds in contrast to a primarily virtual one.

2 Related work

There is already a framework by which user experience can be evaluated, the USUS framework (Weiss, Bernhaupt, Lankes, & Tscheligi, 2009). This framework is divided into four evaluation factors; Usability, Social acceptance, User experience and Societal impact. This paper looks at some elements from the factors Usability and User experience (Tung & Au, 2018, p. 2682–2693).

Efficiency, an aspect of Usability, is defined by the USUS framework as "the rate or speed at which a robot can accurately and successfully assist humans". This paper will investigate if the level of communication and interaction between a robot and a human is correlated with understanding directions.

Utility, an aspect of Usability, is defined by the USUS framework as "how an interface can be used to reach a certain goal or to perform a certain task.".

Embodiment, an aspect of User experience, is defined by the USUS framework as "the relationship between a system and its environment … measured by investigating the different perturbatory channels like morphology". The assumption is often that a humanoid robot will be easier to interact with, as "A robot with physical expressions in the form of a body, or virtually simulated body, can better engage with humans in meaningful interactions". This assumption is often found in literature (Brooks, 2002). A humanoid robot can however also generate unrealistic expectations that a machine with no humanoid features would not have and in extreme cases it can even lead to fear (Tung & Au, 2018, p. 2690).

A previous experiment (Tussyadiah & Park, 2018, p. 11) with a Nao robot, a smaller humanoid robot comparable to Pepper, showed that anthropomorphism was significant in inducing use intention of the Nao robot. People's attention was also focused on the face, with limitation fixation on other body parts, a fact that we expect will change when body parts such as the arms play a more important role, such as when used to point to a certain direction.

Gestures and other hand movements such as pointing will also prevent confusion about which left is meant, the robot's or the human's left. There is a large percentage of the population that has difficulty with distinguishing left from right in itself as according to an Australian study, at least one-third of people at least sometimes experienced frustration with everyday situations that involve the discrimination of left and right (McMonnies, 1990).

Emotions, an aspect of User experience, is defined by the USUS framework as “emotional episodes that are aroused during the interaction with a product [robot]”. These episodes can be divided into fright for the robot, anger as the robot is unable to perform its task sufficiently, satisfaction as the robot is able to perform its task and finally joy if the customer’s expectations are exceeded. As the Pepper robot is a human-shaped robot, expectations turn out to be higher than a robot with a non-human form. Users expect human-like experiences from a highly anthropomorphic robot. Users are more lenient to service failures from a non-humanoid functional robot, with one user clearly differentiating the robot from the responsibilities of humans (Tung & Au, 2018). As hospitality robots are still new, first reactions are likely to be positive as this technology still has a kind of “Wow factor”. Many of the users of that experiment indicated that it was their first encounter with robots, and the sense of uniqueness and novelty combined with a futuristic vibe made some guests rethink what a typical hotel experience could be. Negative reactions were based on the lack of convenience compared to the alternative and that it was just a gimmick.

Human-oriented perception, an aspect of User experience, is defined by the USUS framework as “a robot’s ability to respond to a user’s command as well as detect and interpret events that occurred”. It also states that humans possess powerful sensory capacities that allow them to detect and interpret events that are occurring and respond to them accordingly. Robots that have a high human-oriented perception would be perceived as intelligent. As previously stated in the factor embodiment, a previous experiment with a Nao robot (Tussyadiah & Park, 2018), people paid little attention to some functionality added to better indicate animacy such as hand movement compared to facial expressions. This is in contrast to the fact that it better mimics natural human behaviour.

3 Research question

The research question is ”When requesting directions from a robotic assistant, will there be a difference in user experience and utility if these directions are given visually, verbally, supported by gestures or a combination of them?” Giving directions is an important task for receptionists and as a humanoid robot can more effectively simulate natural human behaviour, there could be a difference in user experience if the directions are given in a more natural way.

3.1 Hypotheses

Our hypothesis is that there is a statistically significant difference based on user experience and utility when a robot gives directions through visual means, auditory means, gestures or a combination of these. A combination of these means better mimics natural human behaviour when someone is explaining how to get somewhere, and would therefore result is a higher score for user experience and utility. It is expected that this effect will be less perceptible due to the existence of a “Wow factor” with this new technology. If every category of user experience and utility is rated higher than would otherwise be warranted, then the possible difference between various conditions is reduced.

4 Method

To answer the research question a Pepper robot was programmed to be able to support the reception personnel with some tasks, namely by greeting people and being able to give directions to any room in the building. The robot is able to operate based on four conditions that either enable or disable functionalities of the robot, such as expressions, gestures and movement to either increase or decrease the intensity of human-robot interaction. Participants will then be able to ask for directions to a given room under these various conditions and then answer questions on their experience. The exact scenario is stated in Section 4.2.

Their experience is assessed through a combination of the Godspeed Questionnaire (Bartneck, Kanda, Mubin, & Mahmud, 2009), a standardised measurement tool for human robot interaction, and a standardised list of questions about Utility.

4.1 Materials

In this investigation a Pepper robot is used, which is able to converse with a user through speech, gestures and a tablet-like interface. This robot is equipped with a wide array of sensors and ways to interact with a user.

The reception interaction program uses code from the Aldebaran library, a library released by Softbank for the development of Naoqi robots, such as Pepper (Softbank Robotics, n.d.).

The Godspeed Questionnaire is used to assess the user experience value of various conditions of human-robot interaction. This questionnaire assesses Anthropomorphism, Animacy, Likeability, Perceived Intelligence and Perceived Safety. In addition we will use some standardised questions about the "Perceived ease of use", "The degree to which the user believes that using the system would be free of effort" and the "Perceived usefulness", "The degree to which a person believes that using the system would enhance his or her daily activities" (Heerink, Krose, Evers, & Wielinga, 2009).

4.2 Experimental setup/approach

The benefits of a physical assistant when needing directions are measured by creating various degrees of interaction. Participants will interact with the robot and ask directions to a given room. The robot will perform based on one of these conditions.

A participant will interact with a robot that operates on a pre-selected condition. The participant can then either try a name or room number from that he knows in the Snellius building or be given a list of names and room numbers to choose from. Through this interaction, the effectiveness of various kinds of human-robot interaction can be measured.

The measurement itself will be done through one of four Google forms which consist of the Godspeed Questionnaire and some questions about utility. Each form corresponds to a single condition. This is done to prevent participants from giving a condition a lower score because it is not the final condition. The questions themselves are the same in each of the four Google forms. The exact question list is included in the appendix.

During the experiment, the participant will be able to interact with the robot using speech or the tablet interface on the robot. This will remain constant during the experiment. What will change however, are the conditions that the robot will perform under.

Every condition consists of the previous condition, expanded through some new kind of interaction. The various conditions are:

(Condition 1 - Tablet + Ambient robot) Directions to the room will be displayed on a tablet.

(Condition 2 - Tablet + Ambient robot + listen to names/room numbers) In addition to interacting with the tablet, participants will be able to voice their request, which the robot will then display on its tablet. It will also have basic ambient behaviour, which means it will be able to greet people and perform some other non-functional behaviour.

(Condition 3 - Tablet + Ambient robot + listen to names/room numbers + voicing directions) In addition to voicing request to the robot, the robot will now be able to give directions through the use of speech.

(Condition 4 - Tablet + Ambient robot + listen to names/room numbers + voicing directions + pointing/looking at directions) In addition to giving directions through speech, the robot will now be able to either point at or look into the direction of the room.

4.3 Measures

The first factor, usability, will be qualitatively measured on an ordinal scale through a standardised list of question about the “Perceived ease of use”, and the “Perceived usefulness”.

Through adjusting the level of interaction and functionality of the robot, it will become easier/more difficult to interact with the robot. As asking for directions is more natural than understanding that you need to use a tablet, it is expected that the “Perceived ease of use” will show some change during the experiment.

As the directions given in the first condition already offer all the information required to go from the starting point to a destination, it is expected that the “Perceived usefulness”, will not vary during the experiment.

The second factor, User experience, will be qualitatively measured on an ordinal scale through the Godspeed Questionnaire. This questionnaire assesses anthropomorphism, animacy, likeability, perceived intelligence and perceived safety. These scores are then mapped to elements from the USUS framework.

The element embodiment is measured through anthropomorphism and animacy. The element emotions are measured through likeability and perceived safety. The element human-oriented perception is measured through perceived intelligence.

While not changing the form of the robot, changing the ways in which the robot and the human will be able to interact with each other could have a large impact on people's perception of the robot. Disabling movement and/or gestures results in a robot more comparable to a virtual assistant. This will create an anthropomorphic robot that will or will not be capable of anthropomorphic movement and as the robot clearly has arms and a facial experience, limiting the functionalities of the robot could impact the embodiment and emotions factor.

Although the embodiment factor could be impacted, it is expected that the embodiment score will still be quite high in every condition, as the robot used in the experiment remains a human-shaped robot, which scores high on this factor.

The existence of the aforementioned "Wow" or novelty factor will likely cause the average value of either the emotional score or all scores to be higher, making it more difficult to see the real value that people attribute to certain factors.

The human-oriented perception factor would also be impacted as a robot incapable of these kinds of movement, while clearly suggesting that it is capable of them, could be considered less intelligent.

As mentioned in the related work, people's attention seems to be focused on the face instead of hand movement. This makes it likely that the last condition of the experiment would not necessarily raise the score compared to previous levels.

5 Technology

The reception robot program consists of two parts that are able to operate independently. The first part is the robot itself and the second part is the web page that shows a map and other information. These parts are able to operate independently to make it easy to create a standalone version of the web page, which could be used by anyone with an internet device. As the web page is built with the size factor of Pepper's tablet in mind, it does require some changes.

These parts are developed in various programming languages:

- Python
- JavaScipt
- HTML
- WAMP

Python is used by the first part, the robot. Both JavaScipt and HTML are used by the webpage on the tablet. HTML is used for the visual aspects and JavaScipt it used to make changed on the webpage, for example to show the destination or change the information about a given room. WAMP is used to both communicate with the robot and to communicate between the tablet and the robot. This way, the robot can pass the requests that it hears to the tablet to display instructions and the tablet can pass request to the robot to voice those instructions.

5.1 Generating directions

The robot has to be able to give directions as this is one of the most important tasks of reception personnel and given the fact that it is equipped with a tablet, could even be more effective than its human counterpart.

5.1.1 Path

Directions are based on a starting and end point. The starting point is fixed during the experiment, as the robot will always stand in front of the reception area. The end point will change based on the request of the user. The given directions should also represent the shortest route to the destination.

To generate these directions, an unidirectional graph consisting of nodes and edges is used. These nodes represent the given rooms, hallways and stairs in the Snellius building of the Leiden university. A visual representation of this can be seen in Figure 1 and 2 . The directions are not hard-coded so that it is easy to change the start point if it ever needs to change without invalidating the current instructions and to make it easy to either make changes to or add to the current graph. It also makes it easy to always choose the shortest path, even if the graph itself needs to change to accommodate temporarily closed areas.

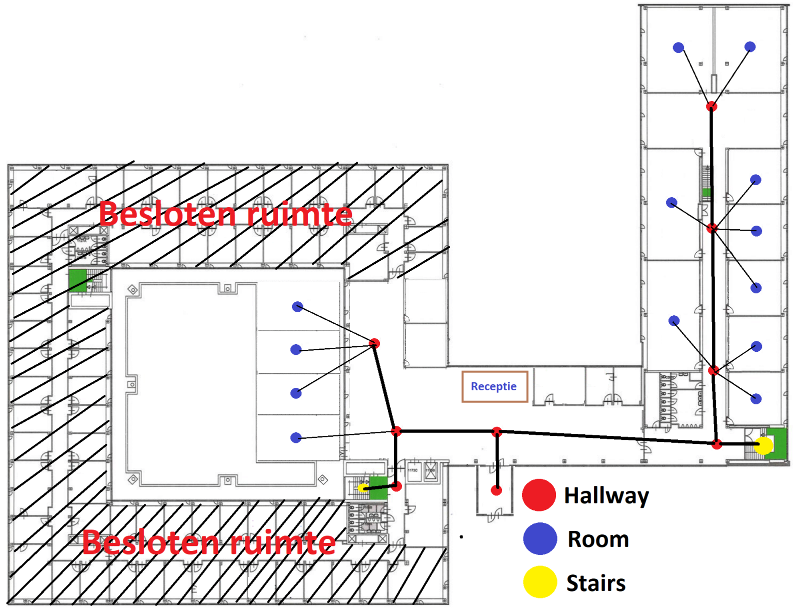


Figure 1: Ground floor graph

Figure 2: First floor graph

A node has an X-coordinate, Y-coordinate, Z-coordinate and a type attribute. The X and Y coordinates are used to determine if a given node is either to the left, to the right or in front of the previous node. The Z-coordinate is used to determine the floor number, which is necessary in case of stairs or elevators. Currently there is no elevator node but this could be added in the future. The type attribute is used to modify the directions text, so that an exception can be made if the node is either the cafeteria, stairs or a room. To generate the directions, we first need to find the shortest path between the starting node and the destination node. This is done through Dijkstra's algorithm. This path is then saved in an array. A line by line explanation of Dijkstra's algorithm used in the experiment can be seen in Listing 3, in the appendix.

5.1.2 Explainable AI

After the path is generated, the location of every node saved in the array is compared to the previous location of the previous one, to determine if the node is either in front of, left of or right of the previous node, as seen in Listing 1. There is also a case for if the current node is behind the previous node but this shouldn't normally happen. As left and right are relative to the direction that the subject moves in, it also requires a heading. As the starting node is always the same in this experiment, the starting node and the direction of the subject is hard-coded in the directions algorithm. For example, the starting direction is always to the north of this map, when the subject needs to go left according to its X and Y coordinates, the direction of the subject changes to west and some text that references that you need to move left through the node type is added to the instruction text. The instructions always end at either a room or the Cafeteria.

```
elif draai == 1: #'gezicht naar oost'
    if dezeGraph.soort[x] == "kamer" or(dezeGraph.soort[x] == "kantine" and nummer + 1 == len(path)):
        if dezeGraph.Ycor[x] - oudY > 20:
            tekst = ". De " + dezeGraph.soort[x] + " bevindt zich aan uw rechterhand."
        elif dezeGraph.Ycor[x] - oudY < 20:
            tekst = ". De " + dezeGraph.soort[x] + " bevindt zich aan uw linkerhand."
        else:
            tekst = ". De " + dezeGraph.soort[x] + " bevindt zich voor u."
    else:
        if dezeGraph.Xcor[x] - oudX > 20:
            if dezeGraph.soort[y] != "gang":
                tekst = "rechtdoor door de " + dezeGraph.soort[y]
            else:
                tekst = "rechtdoor"
        elif dezeGraph.Ycor[x] - oudY < -20:
            tekst = "linksaf door de " + dezeGraph.soort[y]
            draai = 0
        elif dezeGraph.Ycor[x] - oudY > 20:
            tekst = "rechtsaf door de " + dezeGraph.soort[y]
            draai = 2
```

Listing 1: "Explainable AI text generation"

As the result of the solution created by the algorithm, based on the location data and the graph, as easy to understand, this algorithm can be seen as a very simple version of Explainable AI.
As the current method to generate directions as mentioned above results in instruction that are not easy to understand and are clearly generated by an algorithm based on the text, some things need to be improved.

To improve the formatting of the instructions, duplicate instructions that are next to each other are removed, such as in the case of a long hallway. Saying that you need to move through the hallway twice does not add anything and can be combined into one message, as long as there is no possibility to change directions in the hallway. For example, if there is a side hallway to the left, then it is important to again mention that you need to move through the hallway. The loss of clarity while removing instructions is prevented though the type of the node. As a hallway node is not able to have side directions, the node in the hallway would be classified as a corner. This would then prevent the problem as the text is slightly different.

Before mentioning the directions, the text is further improved by adding an opening line, the number of the room that the directions are for and in the case of offices, also naming the person that the room belongs to. This results in the opening text "Om bij kamer " + roomnumber + list of teachers + " te komen gaat u hier ". The list of teachers is also formatted using commas and an " en " for the last name if there is more than one person and it starts with ", het kantoor van ". After this opening text the directions are added according to the algorithm mentioned at the start of this section.

A final improvement is made by having a sentence contain a maximum of three steps, adding punctuation marks and by adding signal words. The steps are separated by step 1 "," step2 "en" step3. After that, a punctuation mark is added followed by the text " Daarna gaat u" and then the next group of steps. To introduce a bit of variety, the first group of steps has the text " Vervolgens gaat u ".

To illustrate the difference that this formatting makes, we present two examples, the first one with formatting and the second one without formatting. The first one is clearly more informative and grammatically more correct.

- Om bij kamer 132, het kantoor van Joost Broekens, te komen gaat u hier linksaf door de gang. Vervolgens gaat u linksaf door de gang, rechtsaf door de gang en de trap op. Daarna gaat u linksaf door de gang, rechtdoor door de kantine en rechtdoor. Daarna gaat u linksaf door de hoek, rechtdoor en linksaf door de hoek. De kamer bevindt zich aan uw rechterhand.

- linksaf door de gang linksaf door de gang rechtsaf door de gang de trap op de trap op linksaf door de gang rechtdoor door de kantine rechtdoor linksaf door de hoek rechtdoor rechtdoor rechtdoor linksaf door de hoek. De kamer bevindt zich aan uw rechterhand.

5.1.3 Generalization

As the instructions themselves do not correspond to a given room number but instead to the type and coordinates of a node, this method of generating directions can be done on any graph that is made using this format. Therefore it is both easy to expand the current graph if the Snellius building would ever be expanded or change the current structure of offices, and the same code can be used on other types of maps. The only requirement is that the user passes on the starting direction and the starting node.

A node consists of the name of the current node, the name of the target node, the distance between the two nodes, the X-coordinate of the target node, the Y-coordinate of the target node, the Z-coordinate of the target node and the type of the target node. In listing 2, the current node is `'Balie'`, the target node is `'Lokalen_Beneden_Gang_Rechts'`, the distance is 60, the X-coordinate is 1700, the Y-coordinate is 800, the Z-coordinate is 0 and the type is `'gang'`.

```

def add_edge_dijk(self, from_node, to_node, weight, Xcor, Ycor, Zcor, soort):

edges = [
 ('Balie', 'Lokalen_Beneden_Gang_Rechts', 60, 1700, 800, 0, 'gang'),
 ]
```

Listing 2: “Adding nodes”

5.2 Pepper robot

The reception program for the Pepper robot is developed in python and uses the Autobahn plugin to connect to with WAMP. The robot responds to speech, touch using the tablet inter connectivity and vision. A picture of the robot can be seen in Figure 3.

When someone stands in front of the robot, it will greet that person and tell him/her what it can do. It is also able to detect and follow the face of that person while he/she is standing in the detection area. Is a face is not detected for 10 seconds, the next time something is detected, the robot assumes that it is a new person and it starts the process again.
The speech part of the robot responds to a list of keywords that we have added. This list consists of:

- Every room number in the Snellius building
- Every room number in the Snellius building with a prefix (“kamer”, “ruimte” or “kantoor”
- Every person with an office in the Snellius building
- Commands to set the condition of the robot
- Non functional commands and trivia that create a more entertaining robot
- Rooms in the Snellius building with a name instead of a number

As the robot has trouble hearing room numbers correctly, the second item adds a list of the room numbers with the prefix that people are most likely to use when asking about a room. This increases the chance that a keyword is matched. Another problem is that the robot is sometimes unable to distinguish between similar room numbers. As we have found no way to solve this problem, this is something that can be improved upon in a further experiment.

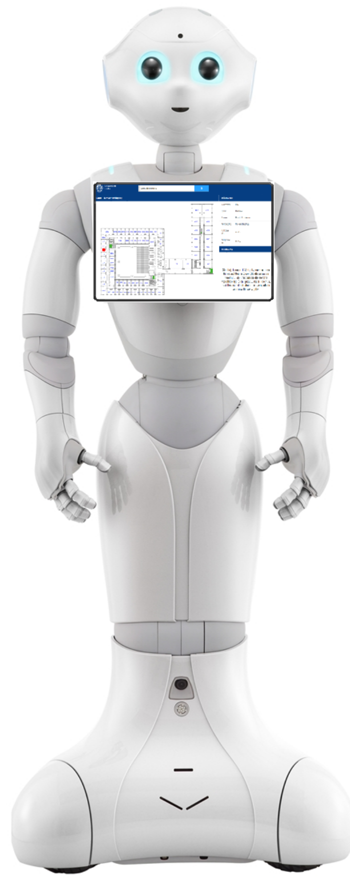


Figure 3: Pepper robot, including the web page

When an approximate match of one of these keywords is heard, this match it first compared with the list of commands to see if a command needs to executed that is not related to giving directions. If this is not the case then the function for giving directions to a room is called, passing along the match as a parameter.

This function first compares the parameter to a list of persons, to see if the parameter is a room number or name or a person. If the parameters matches the name of a teacher then this name is compared to a dictionary of teacher names and room number to get the room number of the office. It then deletes the list of keywords to prevent the robot from giving multiple instructions at once. The room number is then sent to the Dijkstra function that calculates the shortest path, generates instructions, formats them and finally returns the instructions to the previous function.

Based on which condition is active, the robot will also either be able to voice those directions, to point at and look to the first step of the instructions or lead the user there directly. It will also send a remote procedure call (RCP) to the web page with the room number as a parameter, to display it on the tablet.

If someone approaches the robot, it will respond with a form of ambient behaviour. For example, the robot will be able to greet this person or to tell him/her what the robot can do.

5.3 Tablet web page

The HTML web page is written in HTML and uses CSS for the layout. It is then dynamically modified using JavaScript. When a user types in a search term and presses the submit button, this search term is first scanned to see if it is not empty. After that, the search term is compared to a dictionary with all the rooms in the Snellius building, to see if there is a match. Depending on wherever there is a match, one of two things will happen.

If there is a match than another function is called to fill in all the relevant information and another one to highlight the location on the correct map. The first function compares the room number with a dictionary of persons and rooms. If there is a match, the room has to be an office. It will then also fill in the list of persons corresponding to the office in an information field in the top right corner of the page. After determining the type of the room and filling in the person field, it will then search for instructions matching the room to display in the bottom right corner. It will also fill in the floor number. Finally the correct map is shown in the area to the left of the page, highlighting the location of the room. An example of this can be seen in Figure 4. The user either entered 403 or the tablet received a ROP call to display the result for room 403.

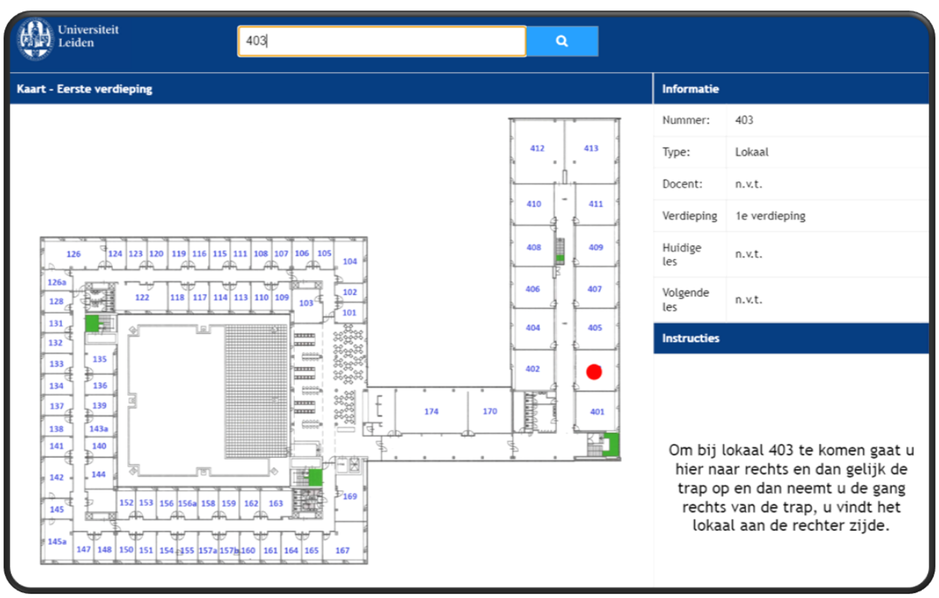


Figure 4: Result of searching for a room number

If there is a WAMP connection, the Pepper robot will also be called, passing the room number as a parameter.

If there is no direct match, the search term is matched with a list of all persons, using 'difflib' to also search for approximate matches. If there is a match, this name is compared to a dictionary of person names and room numbers to get the room number corresponding to that person named. After that the function continues in the same way as if there were a direct match, as seen above. An example of this can be seen in Figure 5. The user entered Broekens, which corresponds to Joost Broekens. This exact scenario cannot be reached through an ROP call as the name would have been translated to a room number already.

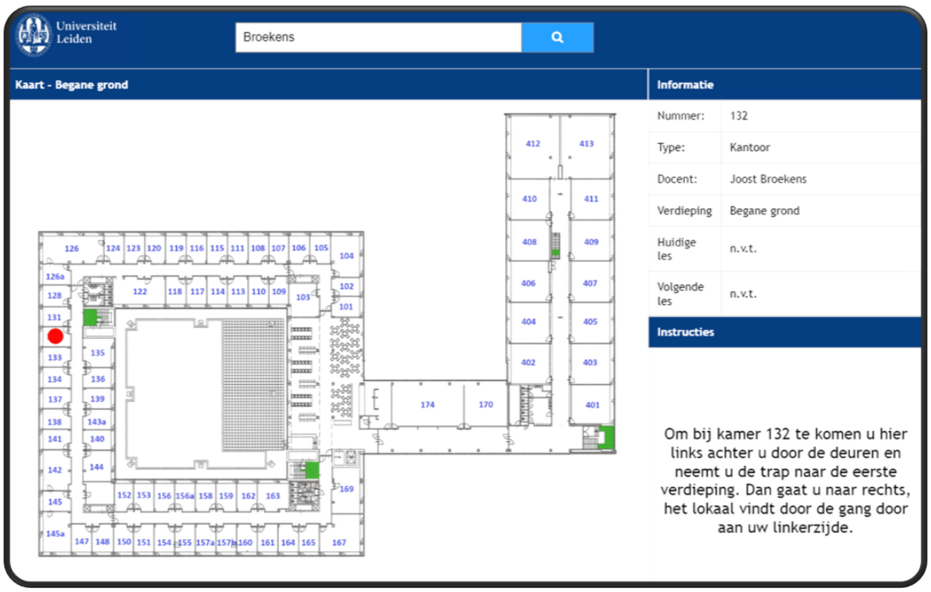


Figure 5: Result of searching for a person's name

If there is no match in any parts of these functions then the web page will return to its default empty state.

5.4 Interconnectivity

The Pepper robot and the web page on the tablet communicate through the WAMP protocol. Both parts connect to the same realm and URL to be register functions. Through a RCP, the tablet or the Pepper robot can than call to this registered function and pass along parameters. An overview of this can be seen in Figure 6.

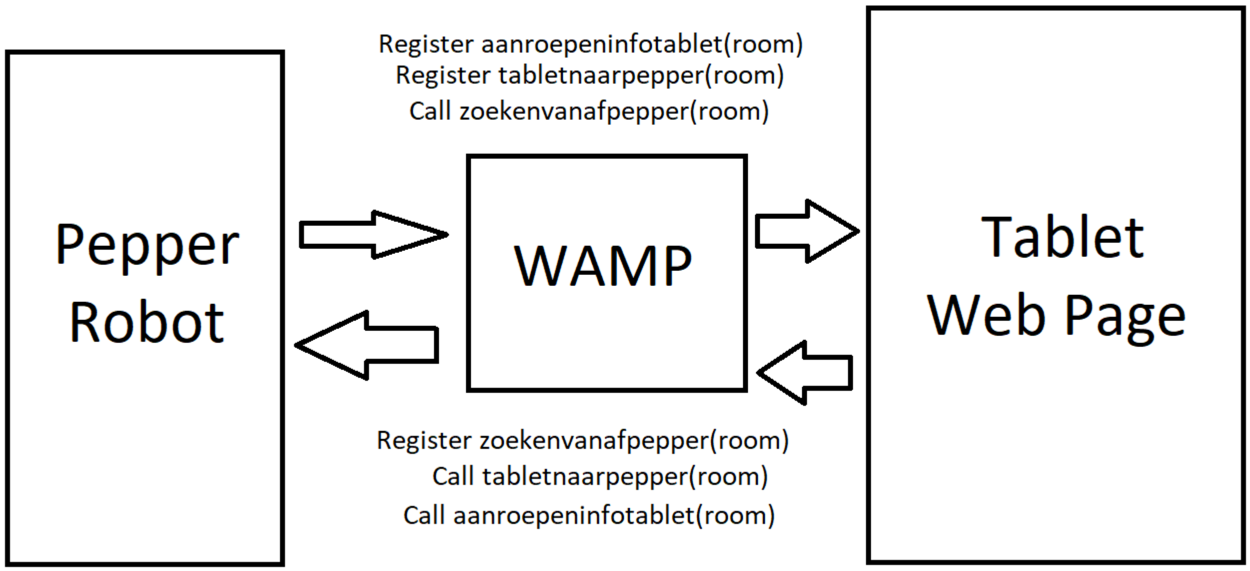


Figure 6: Overview of interconnectivity

WAMP is used in the tablet web page to register a function that receives a room number from the Pepper robot and then displays the location and further information. This is also used in combination with the search button and field to be able to call a function on the Pepper robot while passing the room number as a parameter to make the robot say the instructions corresponding to that room number.

WAMP is used on the Pepper robot to register the previously mentioned function that receives a room number and then gives the corresponding instructions. The robot can also call the web page to pass along the room number for which information has to be given.

While the tablet is able to operate independently of the robot, through a WAMP session it can also receive information generated by the robot. If there is a WAMP session with the robot then the instructions displayed on the tablet are overwritten by the information generated by the explainable AI algorithm. This way, there is always parity between the instructions that the robot gives and the instructions displayed on the tablet. This way, information only has to be updated on the robot instead of on both the tablet and the robot. If there is no WAMP session that it can use the fallback option listed in Section 5.3.

6 Results

In this section, the results of the experiment are listed. An overview of the averages per category can be seen in Figure 7.

Due to some technical problems with the robot during the experiment and the low visitor number during the corona virus, 75% of the results are based on videos showing the various conditions. This also resulted in the need to drop the final condition, as to get enough participants for the other conditions.

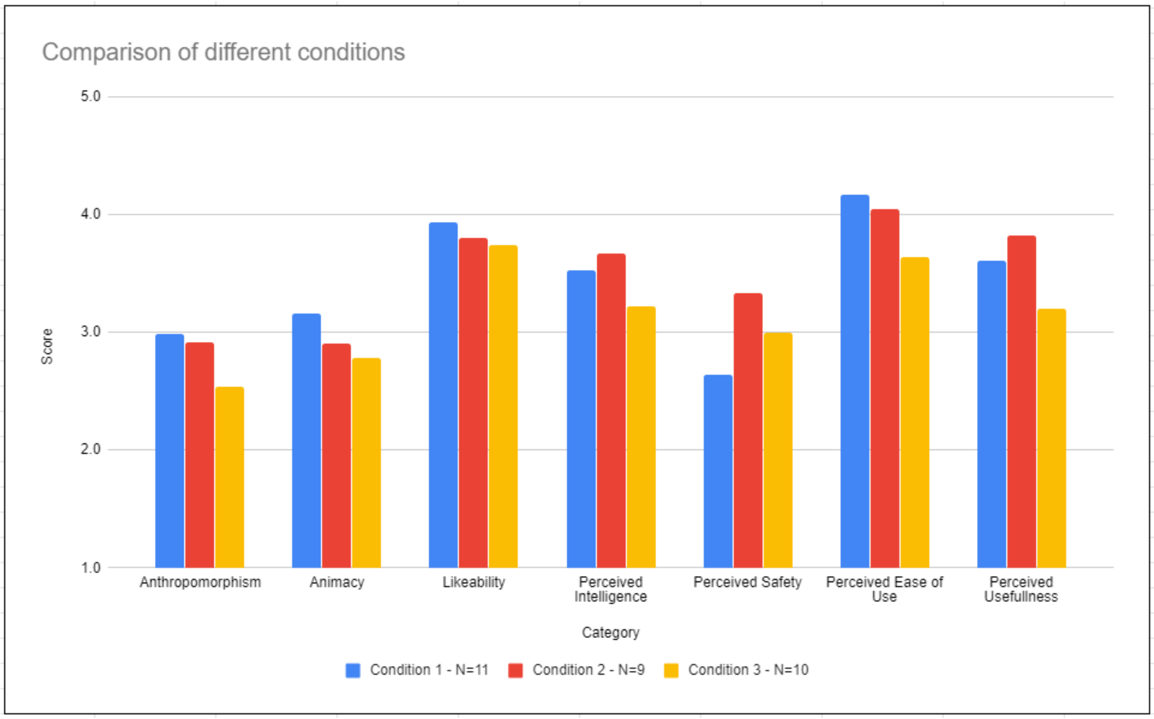


Figure 7: Results of the average score per category overall.

The results are somewhat surprising as the average score seems to decrease in many categories. Only 1 condition, Perceived safety, shows an increase in condition 2 and 3 versus condition 1. Four categories, Anthropomorphism, Animacy, Likeability and Perceived usefulness show a decrease over an increase in functionality.

During the experiment, people over 50 seemed to be more positive about the robot, positively being influenced by the “Wow factor ” about new technology while the younger generation seemed to not be influenced as much. Condition 1 was primarily tested by participants in the Snellius building, which due to the corona virus, were mostly staff and therefore older, which could be the reason that the average score in many categories of condition 1 is higher than other conditions. The other conditions, 2 and 3, where condition 2 was entirely done through video showings, were mostly done by younger participants and this could therefore result in a lower score.

6.1 Participants

Condition 1 was tested by 11 participants. The average age of all participants was 39 years old. In this group there were 6 male participants and 5 female participants.

Condition 2 was tested by 9 participants. The average age of all participants was 38 years. In this group there were 4 male participants and 5 female participants.

Condition 3 was tested by 10 participants. The average age of all participants was 37 years. In this group there were 7 male participants and 3 female participants.

While the average age remained relatively constant, the distribution of ages varied in the participant groups. Condition 1 had 3 participants below 25, 5 middle age participants and 3 participants above 57. Condition 2 had 5 participants below 30 and 4 above 57. Condition 3 had 3 participants aged 21, 5 middle age participants and 2 participants older than 64.

The ratio of male to female participants also varied in condition 3 compared to condition 1 and 2. This can be seen in Figure 8.

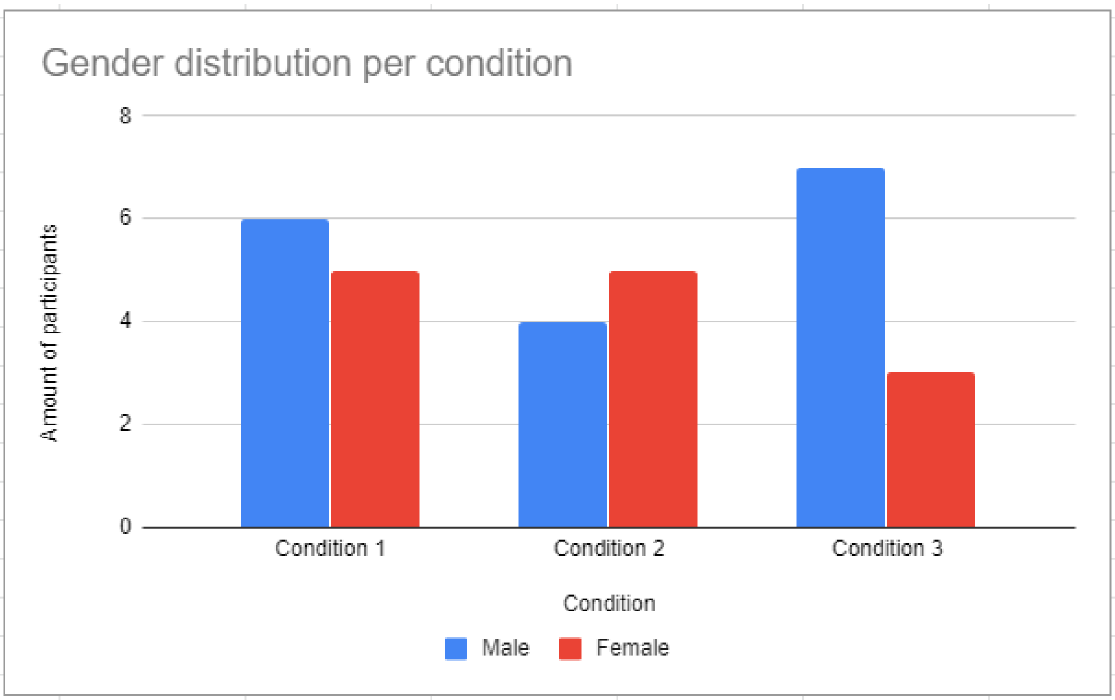


Figure 8: Gender distribution per condition

6.2 Godspeed questionnaire

The Godspeed Questionnaire was used to evaluate 5 categories of user experience. These categories are Anthropomorphism, Animacy, Likeability, Perceived intelligence and Perceived safety

6.2.1 Anthropomorphism

Table 1 contains the descriptive statistics of the three conditions, based on an average of five questions in the category Anthropomorphism.

					95% confidence interval			
	N	Average	Standard deviation	Standard error	lower limit	upper limit	Min	Max
Condition 1	11	2.982	0.8165	0.2462	2.4894	3.4742	1.4	4.8
Condition 2	9	2.911	0.7704	0.2568	2.3975	3.4247	1.2	4.4
Condition 3	10	2.540	0.6720	0.2125	2.1150	2.9650	1.0	4.8
Total	30	2.813	0.7360	0.1344	2.5446	3.0821	1.0	4.8

Table 1: Average score per condition

Table 2 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA ($F(2,27) = 0.6183$, $p = .5463$). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category.

	SS	df	MS	F	P-value	F crit
Between groups	1.145414141	2	0.572707071	0.618294805	0.54631805	3.354130829
Within groups	25.00925253	27	0.926268612			
Totaal	26.15466667	29				

Table 2: ANOVA difference between groups

It is still strange, however, that the average score decreases as the robot behaviour becomes more natural through the increase in conditions, as it is able to hear questions and responds to those questions through speech. The form of the robot does not change, so this element also should result in the same or an higher score as more features are activated. Consciousness should also either remain the same or increase as the robot is now able to behave more human-like. The visual aspect of the robot does not change during the conditions, so this should also remain constant. The motion aspect is the same during the experiment as the final condition was not able to be tested and should therefore also receive the same score.

Although all elements are therefore likely to either receive the same score or show an increase, the results show that the average score decreases.

6.2.2 Animacy

Table 3 contains the descriptive statistics of the three conditions, based on an average of six questions in the category Animacy.

	N	Average	Standard deviation	Standard error	95% confidence interval lower limit	upper limit	Min	Max
Condition 1	11	3.170	0.6694	0.2018	2.7660	3.5734	1.500	4.833
Condition 2	9	2.907	0.8230	0.2743	2.3587	3.4561	1.667	4.500
Condition 3	10	2.783	0.6833	0.2161	2.3512	3.2155	1.333	4.833
Total	30	2.962	0.7375	0.1346	2.6929	3.2315	1.333	4.833

Table 3: Average score per condition

Table 4 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA (F(2,27) = 0.5007, $p = .6116$). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category.

	SS	df	MS	F	P-value	F crit
Between groups	0.82055780022446	2	0.41027890011223	0.500768244344518	0.611583378099256	3.3541308285292
Within groups	22.1210718294052	27	0.819298956644636			
Total	22.9416296296296	29				

Table 4: ANOVA difference between groups

This aspect also shows a decrease in average score, while the condition increased. As a robot that has increased functionality should become more lifelike, real and show more movement, the average score should have increased. It is possible that the addition of voicing instructions generated by an explainable AI makes the robot seem more artificial and could therefore result in a lower score in the organics element. The robot should also have become more interactive and responsive through the added ways in which participants could interact with the robot, resulting in a higher score.

As condition 2 would not have a reduced organics score due to the fact that robot only displays the instructions on the tablet, this can not explain the lower average score compared to condition 1. It is however possible that the lower average score in condition 3 compared to condition 2 can be attributed to this.

6.2.3 Likeability

Table 5 contains the descriptive statistics of the three conditions, based on an average of five questions in the category Likeability.

	N	Average	Standard deviation	Standard error	95% confidence interval lower limit	upper limit	Min	Max
Condition 1	11	3.927	0.3702	0.1116	3.7040	4.1505	3	4.6
Condition 2	9	3.800	0.8000	0.2667	3.2667	4.3333	1.8	5
Condition 3	10	3.650	0.4900	0.1550	3.3401	3.9599	2.0	5
Total	30	3.797	0.5331	0.0973	3.6020	3.9913	1.8	5

Table 5: Average score per condition

Table 6 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA ($F(2,27) = 0.3588$, $p = .7018$). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category.

	SS	df	MS	F	P-value	F crit
Between groups	0.402848484848487	2	0.201424242424244	0.358838721287113	0.701767940098738	3.3541308285292
Within groups	15.1557070707071	27	0.561322484100262			
Total	15.5585555555556	29				

Table 6: ANOVA difference between groups

This aspect should not vary much during the variation of the experiment, as this factor is not related to receiving instructions in a certain way, as the robot is always able to operate on an ambient level, in which it greets the participant and can give other minor information. As there is no reason for the average score to either increase or decrease, this aspect seemed to stay relatively constant.

6.2.4 Perceived intelligence

Table 7 contains the descriptive statistics of the three conditions, based on an average of five questions in the category Perceived Intelligence.

	N	Average	Standard deviation	Standard error	95% confidence interval lower limit	upper limit	Min	Max
Condition 1	11	3.527	0.4430	0.1336	3.2601	3.7944	2.8	4.4
Condition 2	9	3.667	0.6074	0.2025	3.2617	4.0716	2.8	5
Condition 3	10	3.220	0.5000	0.1581	2.9038	3.5362	1.8	5
Total	30	3.467	0.5289	0.0966	3.2735	3.6598	1.8	5

Table 7: Average score per condition

Table 6 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA ($F(2,27) = 1.1221$, $p = .3403$). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category.

	SS	df	MS	F	P-value	F crit
Between groups	1.00884848484849	2	0.504424242424244	1.12206776715899	0.340319769188012	3.3541308285292
Within groups	12.1378181818182	27	0.449548821548822			
Totaal	13.1466666666667	29				

Table 8: ANOVA difference between groups

This aspect showed some expected behaviour. The average score showed an increase from condition 1 to 2 while showing a decrease from condition 1 and 2 to condition 3. As the information presented by the robot is the same during the experiment, the perceived intelligence of the robot should generally remain constant. An explanation for the increase from condition 1 to condition 2 could be that a robot that responds to voice commands seems smarter. The decrease from condition 1 and 2 to condition 3 could be explained through the fact that, while being able to voice the instructions makes the robot seem smarter, the unnatural voice and instructions themselves make the robot seem less intelligent.

6.2.5 Perceived safety

Table 9 contains the descriptive statistics of the three conditions, based on an average of three questions in the category Perceived Safety.

	N	Average	Standard deviation	Standard error	95% confidence interval lower limit	upper limit	Min	Max
Condition 1	11	2.636	0.3967	0.1196	2.3971	2.8756	1.6667	4.0000
Condition 2	9	3.333	0.7407	0.2469	2.8395	3.8272	2.0000	4.6667
Condition 3	10	3.000	0.4000	0.1265	2.7470	3.2530	2.3333	4.6667
Total	30	2.967	0.5444	0.0994	2.7679	3.1655	1.6667	4.6667

Table 9: Average score per condition

Table 10 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA (F(2,27) = 2.7517, p = .0817). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category. This is however the lowest p-value of all categories from the Godspeed Questionnaire. It is also the only category that shows an average score increase in subsequent condition compared to the first condition.

	SS	df	MS	F	P-value	F crit
Between groups	2.42121212121212	2	1.21060606060606	2.75165816326531	0.0817310748986747	3.3541308285292
Within groups	11.8787878787879	27	0.439955106621773			
Totaal	14.3	29				

Table 10: ANOVA difference between groups

As the form of the robot does not change, this aspect should remain fairly constant. The results of this category show that both condition 2 and condition 3 have a higher average score than condition 1. This can be explained in two ways. The first way is that as condition 2 and 3 were mostly tested through video, participants had a different experience with the robot. Participants do not get the same sense of scale of the robot and would therefore feel less scared than if they would physically be next to the robot. The second way is that as the robot had more capabilities, participants were more excited about interacting with the robot.

6.3 User experience

The second part of the questionnaire asks questions about the “Perceived ease of use” and the “Perceived Usefulness”. These categories also show strange results.

As the ways in which a participant can interact with the robot increase, the perceived ease of use should also increase. As no functionality is taken away, the robot should not be seen as more difficult to use. A participant also remarked in condition 2 that he/she would prefer the robot to talk back after a request, further suggesting that the Perceived ease of use should increase instead of decrease.

While the “Perceived ease of use” shows a decrease as conditions increase, the “Perceived usefulness” shows an increase in condition 2, suggesting that the addition of the robot being able to respond based on direction request should increase this factor. Combining this functionality with the robot being able to voice that request should then either have the same or a higher score. This is not the case in this experiment as the average score of Perceived usefulness of condition 3 is even lower than the average score of Perceived usefulness of condition 1.

6.3.1 Perceived ease of use

Table 11 contains the descriptive statistics of the three conditions, based on an average of five questions in the category Perceived ease of use.

					95% confidence interval			
	N	Average	Standard deviation	Standard error	lower limit	upper limit	Min	Max
Condition 1	11	4.164	0.6281	0.1894	3.7849	4.5424	2.6000	5.0000
Condition 2	9	4.044	0.4049	0.1350	3.7745	4.3144	3.0000	5.0000
Condition 3	10	3.640	0.5280	0.1670	3.3061	3.9739	2.8000	5.0000
Total	30	3.953	0.5862	0.1070	3.7393	4.1674	2.6000	5.0000

Table 11: Average score per condition

Table 12 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA ($F(2,27) = 1.8352$, $p = .1797$). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category.

	SS	df	MS	F	P-value	F crit
Between groups	1.5429898989899	2	0.771494949494951	1.53710599754326	0.233232114549404	3.3541308285292
Within groups	13.5516767676768	27	0.501913954358399			
Totaal	15.0946666666667	29				

Table 12: ANOVA difference between groups

This aspect shows a similar average score in condition 1 and 2, with a decrease in condition 3. As every condition builds up on the previous condition, the ease in which participants are able to interact with the robot should increase. Therefore, this score should always either remain constant or increase. A possible explanation about the lower score in condition 3 is that it was not clear through video that one could either use the tablet or voice commands. As this condition then would not build up on a previous condition, it is possible that the score would be lower than previous scores.

6.3.2 Perceived usefulness

Table 13 contains the descriptive statistics of the three conditions, based on an average of three questions in the category "Perceived Usefulness".

					95% confidence interval			
	N	Average	Standard deviation	Standard error	lower limit	upper limit	Min	Max
Condition 1	11	3.867	0.6667	0.2108	3.4450	4.2883	2.0000	5.0000
Condition 2	9	3.815	0.5761	0.1920	3.4307	4.1989	2.0000	5.0000
Condition 3	10	3.200	0.6933	0.2193	2.7615	3.6385	1.6667	5.0000
Total	29	3.621	0.6413	0.1191	3.3825	3.8589	1.6667	5.0000

Table 13: Average score per condition

Table 14 contains the outcome of the one-way between-groups ANOVA test. There were no statistically significant differences between group means as determined by one-way ANOVA (F(2,27) = 1.8352, p = .1797). As there is no significant difference between the means of the three conditions, there is no reason to investigate the significant difference between individual groups in this category.

	SS	df	MS	F	P-value	F crit
Between groups	2.71400595998297	2	1.35700297999148	1.83524805796434	0.179651570308299	3.36901635949544
Within groups	19.2246913580247	27	0.739411206077873			
Totaal	21.9386973180077	29				

Table 14: ANOVA difference between groups

This aspect also shows a similar average score in condition 1 and 2, with a decrease in condition 3. As every condition includes all functionality of the previous condition, the usefulness of the robot score should either remain constant or increase. If the tablet already has a certain level of usefulness, then being able to interact with the robot by voice should not negatively impact this and could for some people even be seen as an upgrade. A possible explanation about the lower score in condition 3 is that it was not clear through video that one could either use the tablet or voice commands. As this condition then would not build up on a previous condition, it is possible that the score would be lower than previous scores. During the experiment, there were also problems with keyword matching, possibly negatively impacting the score of people that did interact with the robot in real life.

6.4 Additional remarks

The hands on experiment and the additional feedback from the final question of the questionnaire, resulted in further remarks outside of a quantifiable score.
5 out of 30 participants mentioned that they thought that the robot is cool, one participant even mentioned that these kinds of robots are likely the future of hospitality interaction.

There were also two participants that would prefer a human receptionist instead of a robot, as how great the robot could be, it would still remain a “dead object”. One person even remarked that he would “prefer to be helped by an attractive woman than a robot, even if he would be shown the wrong room”.

Participants also gave remarks about the voice of the robot. Two participants remarked that the voice of the robot could either be lower and/or less fast and that they would prefer a more natural voice.

There were also some assumptions made about the tablet interface that were not part of the experiment. For example, most participants expected that the enter button would work the same as the search button. So when using the search bar, after entering the search term the enter button should immediately start the function to present the directions. In the current experiment the enter button did nothing and some people needed to be helped to press the looking glass symbol next to the search bar. One participant also expected to be able to press on a room on the map, instead of using the search bar.

Many older participants also had some difficulty with the questionnaire. While opening the questionnaire was easy, they found it difficult to choose a number for every question. Additional information at each question could aid them in understanding what each question meant.

7 Conclusions

This experiment tried to determine the effect of supporting directions by auditory means, visual means, gestures or a combination of them when receiving these directions from a robot. Based on three conditions, as the final condition was unable to be tested, the average score per category of the Godspeed Questionnaire and additional question about utility was calculated and the effect of these various means was measured.

The results of this experiment show that there is no statistically significant difference between any of the three conditions in any of the seven categories. This means that the earlier stated hypothesis is incorrect. There is no statistically significant difference based on user experience and utility when a robot gives directions through visual means, auditory means, gestures or a combination of these.

While the experiment did not prove the hypothesis, it did result in some useful feedback. Based on participants’ reactions to the robot, a better solution could be to have a mix of both robots and humans at the reception desk. This way, people have a choice in interacting with either a human or a robot. The voice and unnatural sounding instructions could also have impacted the score of the experiment. Many people also make assumptions about how to interact with a web page on a tablet, which should be built into the experiment. Finally the questionnaire itself could be unclear for participants, so it is necessary to properly explain every question instead of only providing a category description.

As mentioned in Section 6, the resulting scores could also be impacted through the use of video segments, based on which participants filled in the questionnaire. The low sample size of the experiment could also mean that the results of this experiment do not accurately reflect reality.

To determine if this is the case, this or a variation of this experiment needs to be done in the future. It is recommended that this is done after the measures needed to be taken to minimize the effect of the corona virus are lifted, ensuring that there is enough foot traffic in buildings to gather a large number of responses, that are not gathered through video segments but real life experience.

8 Further research

Based on the current experiment, some improvements can be made to both expand the scope of the experiment and to improve the current setup of the experiment.

8.1 Improvements upon the current experiment

The current experiment was impacted by both the corona virus and technical issues during the testing phase.

Due to the low visitor number of the experiment building, as only a part of the staff and almost none of the students where allowed in the building, the last condition of the experiment could not be tested. It is still interesting, however, if there is a statistically relevant effect in user experience and or utility if the robot is able to point at or look into the direction that people need to go to. Therefore further research should be done that also investigates this condition.

The experiment could also benefit from more hands on feedback as the robot currently needed to be reset after every use due to some function call crashing the robot. Due to this crashing, it was required to do more video testing which results in more responses but less feedback about how a specific participant would use the robot. For example, it turned out that when including a tablet, people expect to be able to press on the room in the map to also receive instruction to that location.

There were also some technical issues relating to the hearing of keywords, resulting in the robot either hearing an incorrect name of an incorrect room number. Due to the amount of keywords listed on the robot during the experiment and the similarity of the room number, either most or all speech to text software would have difficulty in recognising the correct keyword.

This results in the robot giving directions to the wrong room number. The participant will then either have to repeat the request or will have to use the tablet to bypass this recognition problem. As an investigation about user experience does not require the robot to be functional for every given room in the building, only the ones that are being tested, the other keywords related to rooms not used in the experiment could be influenced without influencing any functionality used in the test. By influencing these keywords, the possibility for an incorrect match is reduced, leading to a more positive user experience. Another way to improve the accuracy of hearing keywords could be by using different speech-to-text software, for example by using Google's solution for transforming audio into text.

8.2 Expanding the scope

As this experiment only looked at four conditions and was able to gather data about only three of those, a way to expand the scope would be by adding more conditions.

For example, one condition could be by testing only the tablet without the presence of a robot. There is no utility difference between this new condition and condition 1. There is also no difference between this new condition and condition 2 of the current experiment, given that the robot is able to hear and understand voice commands of the user.

Another condition that could be tested is by having the robot be able to move into the direction that the participant needs to go to. Given that the Pepper robot itself is quite slow, this condition could either be seen as an improvement or only hindrance. To properly implement this condition the robot needs to know its current position, the position of the destination and how to get to its destination. One way to implement this in the current experiment is to use two large distinct markers that function as the destination point for the robot. By keeping this marker in the center of the robot focal point, the size of the object can be measured. Comparing this size to the real size of the object allows the robot to estimate the distance to that object. The robot also has to keep adjusting its direction to keep this object in the middle of its focal point. Using this the robot can find its way to the destination point. When the robot has to return to its base point, it can use the other marker to first get the direction to this point and then by calculating the distance from the robot to that point, how far it is away from the base point. If the robot has reached the base point then all it has to do is rotate so that it faces the front door.

Another condition that could be tested is the influence of the specific voice and the number of words per minute of the robot. One participant remarked that the robot talks too fast and that he/she would prefer a more natural voice. To test this influence, one can vary the words per minute of the speech part of the robot or change the voice of the robot. An easy way to implement this is to use Google's text-to-speech API. These voices resemble natural speech and give the option to choose from a number of both male and female voices, to control the pitch and to control the the number of words per minute.

There is also the possibility that the results of this experiment where influenced by the type of participant interacting with the robot. While the questionnaire used in the experiment lacked an age field, there seemed to be a difference between the age of the participant and the score that the participant gave according to the Godspeed test. People over 50 seemed to be more impressed by the technology and therefore gave the robot a higher score in most categories. This effect could possibly be related to the "Wow factor" which would lead to higher scores in all conditions. Younger participants where less positive about the robot and therefore gave it a lower score. Further research could determine if there really is a correlation between age and user experience and utility. Other possible factors could be occupation or educational level.

References

Bartneck, C., Kanda, T., Mubin, O., & Mahmud, A. (2009). Does the design of a robot influence its animacy and perceived intelligence? *International Journal of Social Robotics*, *1*, 195-204. doi: 10.1007/s12369-009-0013-7

Brooks, R. (2002). Humanoid robots. *Commun. ACM*, *45*(3), 33–38. doi: 10.1145/504729.504751

Ebling, M. (2016). Can cognitive assistants disappear? *IEEE Pervasive Computing*, *15*, 4-6. doi: 10.1109/MPRV.2016.41

Foster, M. (2018). *Aging japan: Robots may have role in future of elder care.* Thomson Reuters. Retrieved July 09, 2020, from https://www.reuters.com/article/us-japan-ageing-robots-widerimage/aging-japan-robots-may-have-role-in-future-of-elder-care-idUSKBN1H33AB

Heerink, M., Krose, B., Evers, V., & Wielinga, B. (2009, 11). Measuring acceptance of an assistive social robot: A suggested toolkit. In (p. 528 - 533). doi: 10.1109/ROMAN.2009.5326320

Herron, J. (2017). Intelligent agents for the library". *Journal of Electronic Resources in Medical Libraries. Routledge*, *14*(3–4), 139. doi: 10.1080/15424065.2017.1367633.

Lagrue, E., & Busy, M. (2020). *Caresses: Ssmart and friendly robots for the elderly.* Softbank Robotics. Retrieved July 09, 2020, from https://developer.softbankrobotics.com/blog/caresses-smart-and-friendly-robots-elderly

Ly, C. (2016, May). *Mastercard unveils the first commerce application for humanoid robot pepper.* Retrieved July 09, 2020, from https://www.futuretimeline.net/blog/2016/05/30.htm

Matney, L. (2019). *More than 100 million alexa devices have been sold.* Retrieved February 3, 2019, from https://techcrunch.com/2019/01/04/more-than-100-million-alexa-devices-have-been-sold/

McMonnies, C. W. (1990). Left-right discrimination in adults. *Clinical and Experimental Optometry*, *73*(5), 155-158. doi: 10.1111/j.1444-0938.1990.tb03116.x

Mogg, T. (2015). *Softbank's pepper robot sells out in just 60 seconds.* Retrieved February 09, 2020, from https://www.digitaltrends.com/cool-tech/softbanks-emotional-pepper-robot-sells-out-in-just-60-seconds/

Olson, P. (2018). *Softbank's robotics business prepares to scale up.* Retrieved February 09, 2020, from https://www.forbes.com/sites/parmyolson/2018/05/30/softbank-robotics-business-pepper-boston-dynamics/#300d6cd54b7f

Softbank Robotics. (n.d.). *Aldebaran documentation.* Retrieved February 12, 2020, from http://doc.aldebaran.com/2-4/index_dev_guide.html

Tung, V., & Au, N. (2018). Exploring customer experiences with robotics in hospitality". *International Journal of Contemporary Hospitality Management*, *30*(7), 2680–2697. doi: 10.1108/IJCHM-06-2017-0322

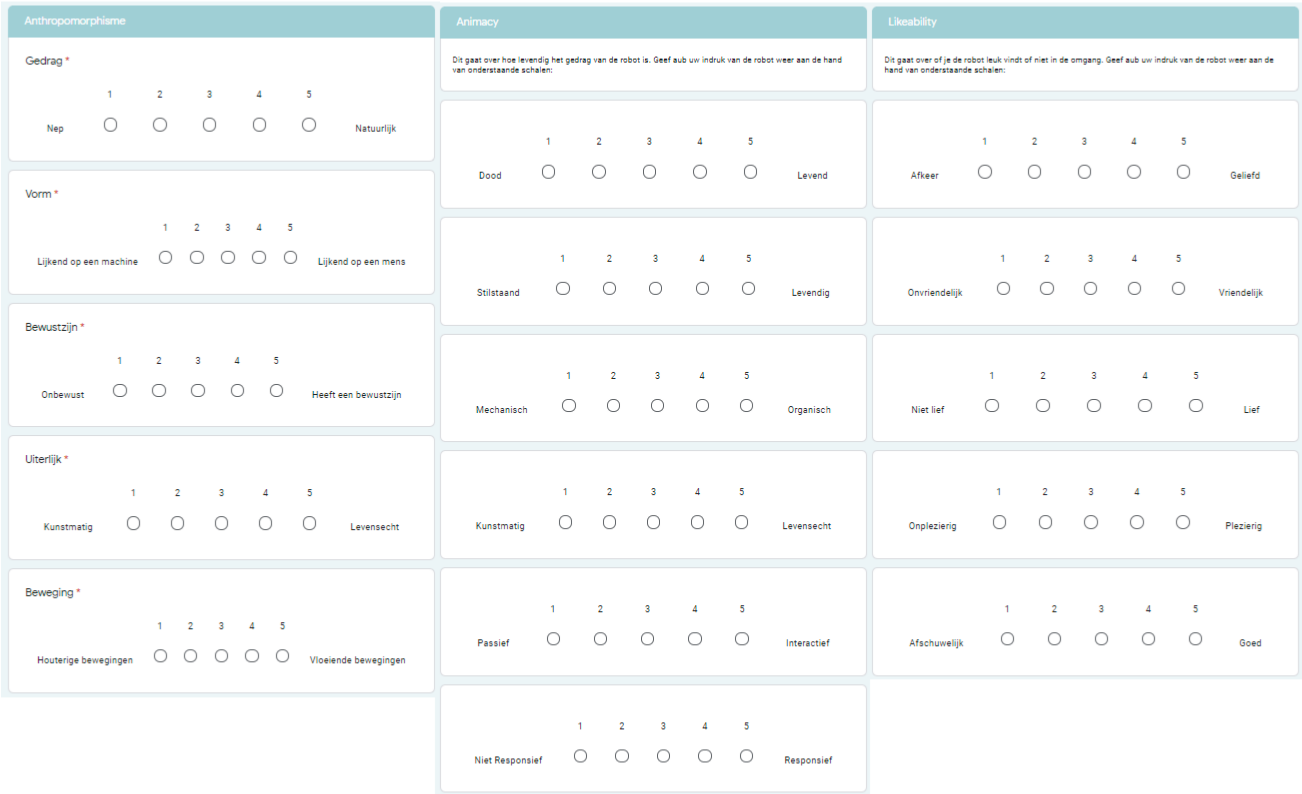
Tussyadiah, I., & Park, S. (2018). *Consumer evaluation of hotel service robots.*

Weiss, A., Bernhaupt, R., Lankes, M., & Tscheligi, M. (2009, 01). The usus evaluation framework for human-robot interaction. *Proc. of AISB 09*, *4*, 11-26.

Appendix

```python
def dijsktra(dezeGraph, initial, end, draaihoek):
    # shortest paths is a dict of nodes
    # whose value is a tuple of (previous node, weight)
    shortest_paths = {initial: (None, 0)} #at the start the weight to every node is none
    current_node = initial
    visited = set() # set of visited nodes

    while current_node != end:
        visited.add(current_node) #marks the current node as visited
        destinations = dezeGraph.edges[current_node] # every node that can be reached from this node
        weight_to_current_node = shortest_paths[current_node][1] #the weight of the current node is the shortest path to that node

        for next_node in destinations: # for each neighbor of the current node
            weight = dezeGraph.weights[(current_node, next_node)] + weight_to_current_node #weight = current + distance to next
            if next_node not in shortest_paths: #if there is no path yet then is is always the shortest path
                shortest_paths[next_node] = (current_node, weight) #shortest path to next node is from the current node with weight
            else: #if there is already a path
                current_shortest_weight = shortest_paths[next_node][1] #get the current shortest path to the next node
                if current_shortest_weight > weight: #if the new weight is less then the current shortest path then
                    shortest_paths[next_node] = (current_node, weight) # shortest path to next node is from the current node with weight

        next_destinations = {node: shortest_paths[node] for node in shortest_paths if node not in visited}
        if not next_destinations: #the room can not be reached
            return "Kamer " + end + " zit niet in dit gebouw, bedoelt u een andere kamer?"
        # next node is the destination with the lowest weight
        current_node = min(next_destinations, key=lambda k: next_destinations[k][1])

    # Work back through destinations in shortest path
    path = []
    while current_node is not None:
        path.append(current_node)
        next_node = shortest_paths[current_node][0]
        current_node = next_node
```

Listing 3: ”Dijkstra’s algorithm”



Anthropomorphisme

		1	2	3	4	5	
Gedrag *	Nep	○	○	○	○	○	Natuurlijk
Vorm *	Lijkend op een machine	○	○	○	○	○	Lijkend op een mens
Bewustzijn *	Onbewust	○	○	○	○	○	Heeft een bewustzijn
Uiterlijk *	Kunstmatig	○	○	○	○	○	Levensecht
Beweging *	Houterige bewegingen	○	○	○	○	○	Vloeiende bewegingen

Animacy

Dit gaat over hoe levendig het gedrag van de robot is. Geef aub uw indruk van de robot weer aan de hand van onderstaande schalen:

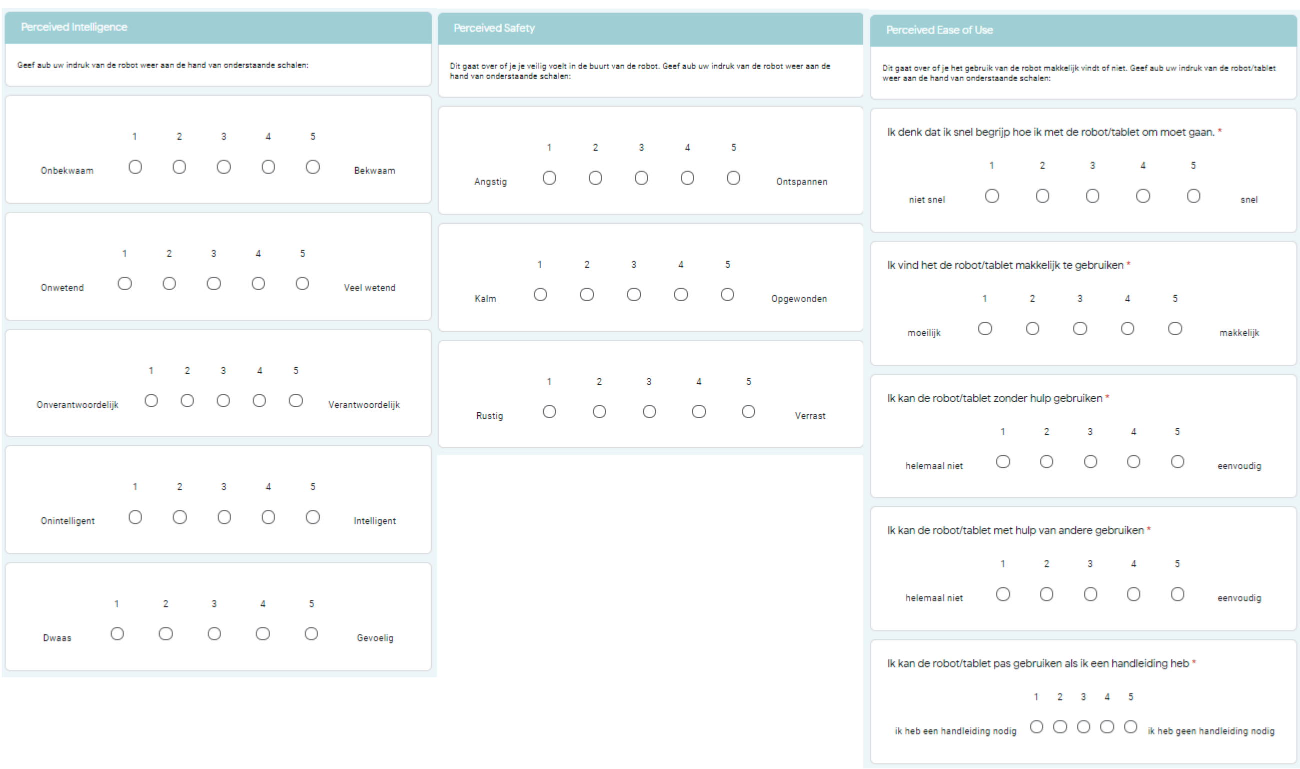
	1	2	3	4	5	
Dood	○	○	○	○	○	Levend
Stilstaand	○	○	○	○	○	Levendig
Mechanisch	○	○	○	○	○	Organisch
Kunstmatig	○	○	○	○	○	Levensecht
Passief	○	○	○	○	○	Interactief
Niet Responsief	○	○	○	○	○	Responsief

Likeability

Dit gaat over of je de robot leuk vindt of niet in de omgang. Geef aub uw indruk van de robot weer aan de hand van onderstaande schalen:

	1	2	3	4	5	
Afkeer	○	○	○	○	○	Geliefd
Onvriendelijk	○	○	○	○	○	Vriendelijk
Niet lief	○	○	○	○	○	Lief
Onplezierig	○	○	○	○	○	Plezierig
Afschuwelijk	○	○	○	○	○	Goed

Figure 9: Questionnaire part 1



Perceived Intelligence

Geef aub uw indruk van de robot weer aan de hand van onderstaande schalen:

	1	2	3	4	5	
Onbekwaam	○	○	○	○	○	Bekwaam
Onwetend	○	○	○	○	○	Veel wetend
Onverantwoordelijk	○	○	○	○	○	Verantwoordelijk
Onintelligent	○	○	○	○	○	Intelligent
Dwaas	○	○	○	○	○	Gevoelig

Perceived Safety

Dit gaat over of je je veilig voelt in de buurt van de robot. Geef aub uw indruk van de robot weer aan de hand van onderstaande schalen:

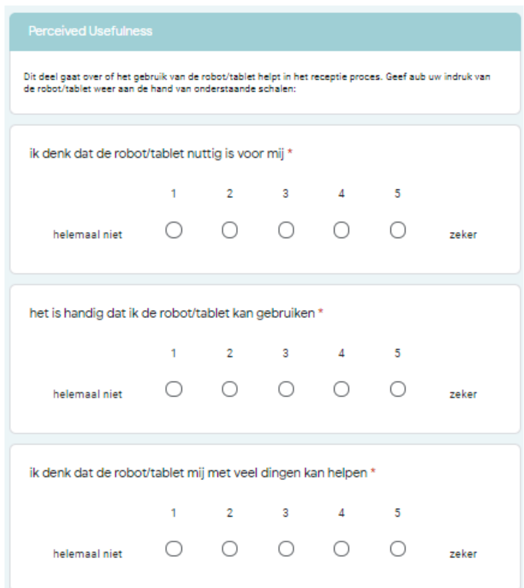
	1	2	3	4	5	
Angstig	○	○	○	○	○	Ontspannen
Kalm	○	○	○	○	○	Opgewonden
Rustig	○	○	○	○	○	Verrast

Perceived Ease of Use

Dit gaat over of je het gebruik van de robot makkelijk vindt of niet. Geef aub uw indruk van de robot/tablet weer aan de hand van onderstaande schalen:

		1	2	3	4	5	
Ik denk dat ik snel begrijp hoe ik met de robot/tablet om moet gaan. *	niet snel	○	○	○	○	○	snel
Ik vind het de robot/tablet makkelijk te gebruiken *	moeilijk	○	○	○	○	○	makkelijk
Ik kan de robot/tablet zonder hulp gebruiken *	helemaal niet	○	○	○	○	○	eenvoudig
Ik kan de robot/tablet met hulp van andere gebruiken *	helemaal niet	○	○	○	○	○	eenvoudig
Ik kan de robot/tablet pas gebruiken als ik een handleiding heb *	ik heb een handleiding nodig	○	○	○	○	○	ik heb geen handleiding nodig

Figure 10: Questionnaire part 2


Perceived Usefulness

Dit deel gaat over of het gebruik van de robot/tablet helpt in het receptie proces. Geef aub uw indruk van de robot/tablet weer aan de hand van onderstaande schalen:

ik denk dat de robot/tablet nuttig is voor mij *

	1	2	3	4	5	
helemaal niet	○	○	○	○	○	zeker

het is handig dat ik de robot/tablet kan gebruiken *

	1	2	3	4	5	
helemaal niet	○	○	○	○	○	zeker

ik denk dat de robot/tablet mij met veel dingen kan helpen *

	1	2	3	4	5	
helemaal niet	○	○	○	○	○	zeker

Figure 11: Questionnaire part 3